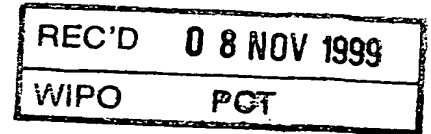


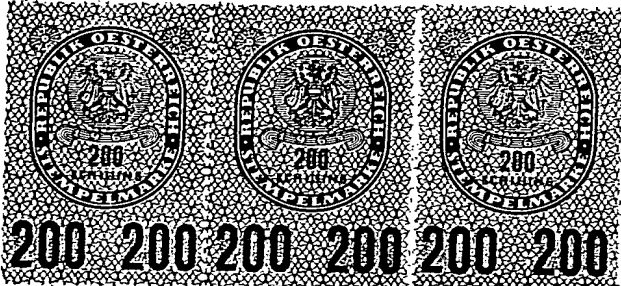
EP
6842

ÖSTERREICHISCHES PATENTAMT

A-1014 WIEN, KOHLMARKT 8 - 10

EU

Aktenzeichen A 1623/98



Das Österreichische Patentamt bestätigt, dass

**die Firma Siemens Aktiengesellschaft Österreich
in A-1210 Wien, Siemensstraße 92,**

am 30. September 1998 eine Patentanmeldung betreffend

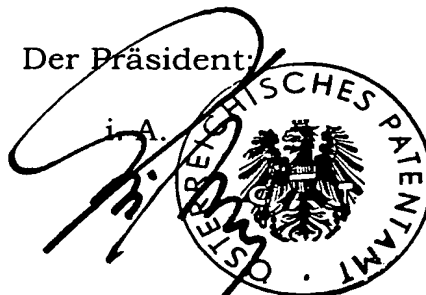
**"Verfahren zur Bestimmung der Lage von Textzeilen bei
Mustererkennungsaufgaben",**

überreicht hat und dass die beigeheftete Beschreibung samt Zeichnung mit der ursprünglichen, zugleich mit dieser Patentanmeldung überreichten Beschreibung samt Zeichnung übereinstimmt.

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH
RULE 17.1(a) OR (b)

Österreichisches Patentamt
Wien, am 25. August 1999

Der Präsident:



Kanzleirat FÜHRLINGER
Fachoberinspektor





ÖSTERREICHISCHES PATENTAMT
Verwaltungsstellen-Direktion

.....240..... s €
Kanzleigegebühr bezahlt.

Balkau

A 1623

98-1

Urtext

Siemens Aktiengesellschaft Österreich

Pm

Wien, 30.09.1998

GR 98 P 7016 AT

Verfahren zur Bestimmung der Lage von Textzeilen bei
Texterkennungsaufgaben

Die Erfindung betrifft ein Verfahren zur Bestimmung der Lage von Textzeilen bei Texterkennungsaufgaben.

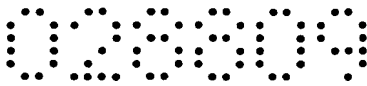
Bei der automatischen Erkennung von Texten, d.h. bei der Umwandlung der grafischen Information eines Dokumentes in Textzeichen, die mittels elektronischen Textverarbeitungsprogrammen weiterverarbeitet werden können, ist eine wesentliche Voraussetzung für einen erfolgreichen Erkennungsvorgang die genaue Bestimmung der Lage und der Größe der einzelnen Zeichen. Voraussetzung dafür ist wiederum die Kenntnis der Lage und der Maße der Textzeilen.

Bei handgeführten Lesevorrichtungen kommt es überdies zu einem nichtlinearen Verlauf der Textzeilen in dem erfaßten Bildausschnitt. Hier besteht die Notwendigkeit, den Verlauf einer Textzeile zu bestimmen.

Der Erfindung liegt daher die Aufgabe zugrunde, ein Verfahren anzugeben, mit dem Lage und Verlauf von Textzeilen bestimmt werden können.

Dies geschieht erfindungsgemäß mit einem Verfahren der eingangs genannten Art, bei dem die Helligkeitsverteilung eines erfaßten Bildausschnittes entlang der Vertikale ermittelt wird und diese Helligkeitsverteilung gefiltert wird, bei dem Maximumwert und Minimumwert der solcherart erhaltenen Funktion ermittelt werden und auf der Grundlage dieser Extrema Schwellwerte berechnet werden, die als Grundlage für die Unterscheidung zwischen Textzeile und Zeilenzwischenraum dienen. Das erfindungsgemäße Verfahren

1/1



ermöglicht die Feststellung der Zeilenlage und Höhe mit geringem Aufwand und damit sehr schnell, sodaß diese Information unmittelbar erhalten wird wodurch ein schneller Erkennungsvorgang ermöglicht wird. Dies ist besonders bei Lesevorrichtungen von Bedeutung, die nur begrenzte elektronische Speicher zur Verfügung haben und deshalb eine rasche Umsetzung der speicherintensiven Bildinformation in Textinformation noch während der Bildaufnahme erfordern.

Vorteilhaft ist eine Ausgestaltung des Verfahrens in der Weise, daß ein Zeilenzwischenraum dann festgestellt wird, wenn die Funktion eine Kombination eines Maximums mit einem Minimum aufweist, bei der das Minimum einen Wert kleiner als $\text{Funktionsminimum} + \text{Anzahl der Bildpunkte über die Breite des Bildausschnittes}/15 + 2 \cdot \text{Anzahl der Bildpunkte über die Breite des Bildausschnittes}/15 \cdot \text{Funktionsmaximum}/\text{Anzahl der Bildpunkte über die Breite des Bildausschnittes}$ aufweist und der Abfall der Funktionswerte nach dem Maximum einen Wert größer $(\text{Funktionsmaximum} - \text{Funktionsminimum})/2$ aufweist. Diese Ausgestaltung hat sich in der Praxis durch sehr gute Ergebnisse bewährt.

Ebenfalls vorteilhaft ist eine weitere Ausgestaltung des Verfahrens in der Weise, daß zur Feststellung des linken Rands der Textzeilen vor der Bestimmung der Lage der Textzeilen die Helligkeitsverteilung eines erfaßten Bildausschnittes entlang der Horizontale ermittelt wird und die solcherart erhaltene Funktion den Zeilenbeginn durch einen sprunghaften Anstieg des Funktionswertes darstellt. Damit ist der Zeilenbeginn auf einfache Weise mit geringem Aufwand zu ermitteln. Weiterhin kann für die Ermittlung der Lage der Textzeilen gewährleistet werden, daß dabei nur Bilder berücksichtigt werden, die tatsächlich Textzeilen enthalten und ein Benutzerfehler wie z.B. das Ansetzen des Lesestiftes zu weit links von dem Zeilenbeginn die Zeilenermittlung nicht beeinflußt.

Günstig ist es, wenn nach der erstmaligen Feststellung der Lage einer Zeile der weitere Verlauf derselben durch Auswertung der Information über die erkannten Textzeichen ermittelt wird. Durch die Auswertung des Ergebnisses der Zeichenklassifikation ist eine besonders genaue Bestimmung des Zeilenverlaufs möglich.

Die Erfindung wird anhand von 2 Figuren näher erläutert. Es zeigen beispielhaft:

Fig. 1 einen Textausschnitt wie er typisch von einer handgeführten Lesevorrichtung erfaßt wird, sowie das daraus ermittelte Hystogramm und

Fig. 2 das gefilterte Hystogramm mit den eingetragenen Parametern zur Beurteilung des Bildes.

Der Ablauf des erfindungsgemäßen Verfahrens ist wie folgt: Zu dem erfaßten Bildausschnitt UB wird ein Zeilenhistogramm $f(y)$ ermittelt. Dabei werden zu jeder Zeile die Werte aller Bildpunkte dieser Zeile (0 für Weiß und 1 für Schwarz) aufsummiert. Ergebnis ist eine Funktion $f(y)$ mit

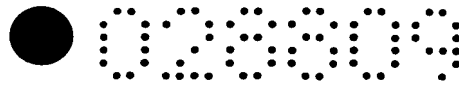
$$f(y) = \sum_{i=0}^{Width-1} (BlackPixel) \quad \text{wobei}$$

y den Zeilenindex des Bildes bedeutet
Width Breite (Spaltenanzahl) des Bildausschnittes angibt.

Diese Funktion hat bei Vorliegen eines Textes einen typischen Verlauf wie er beispielhaft in Fig. 1 dargestellt ist.

In einem weiteren Schritt erfolgt eine Filterung gemäß

$$f'(y) = \frac{\sum_{i=-5}^{+5} (f(y+i) * G(i))}{\sum_{i=-5}^{+5} G(i)}$$



mit

y Index im Zeilenhistogramm
G Gewichtung entsprechend einer exponentiellen
 Glättungskurve
i Index der Glättungskurve

Während des Filtervorganges werden auch der absolute Maximumwert *Valuemax* d.h. die Anzahl der schwarzen Bildpunkte (Black Pixel) der dunkelsten Zeile und das absolute Minimum *Valuemin* d. h. die Anzahl der schwarzen Bildpunkte (Black Pixel) der hellsten Zeile ermittelt.

Aus diesen beiden Werten werden Parameter für die Beurteilung des Bildes abgeleitet. Diese sind:

$Grubenlimit = (Valuemax - Valuemin)/2$
mindestens aber Anzahl der Bildpunkte über die Breite des Bildausschnittes/30

$Minimarand = Valuemin + \text{Anzahl der Bildpunkte über die Breite des Bildausschnittes}/15$
höchstens aber $2 \cdot \text{Anzahl der Bildpunkte über die Breite des Bildausschnittes}/15$

$Minimaschwelle = Minimarand + (2 \cdot \text{Anzahl der Bildpunkte über die Breite des Bildausschnittes}/15 * (Valuemax/\text{Anzahl der Bildpunkte über die Breite des Bildausschnittes}))$
höchstens aber $3 \cdot \text{Anzahl der Bildpunkte über die Breite des Bildausschnittes}/15$

Anhand der Funktion $f'(y)$ und der ermittelten Schwellwerte wie sie beispielhaft in Fig. 2 dargestellt sind, wird nun das erfaßte Bild UB hinsichtlich des Vorliegens von Textzeilen und Zeilenzwischenräumen beurteilt.

Dazu wird der Kurvenverlauf darauf untersucht, ob Werte vorliegen, die geringer als der Parameter *Minimaschwelle*

sind. Ist dies der Fall, dann wird der betroffene Bereich als gültiges Minimum und damit als möglicher Zeilenzwischenraum qualifiziert.

Ein tatsächlicher Zeilenzwischenraum liegt jedoch nur dann vor, wenn durch ein angrenzendes Maximum mit einer gewissen Ausprägung das Vorhandensein einer Textzeile angezeigt wird. Diese gültigen Maxima sind durch ein anschließendes Absinken des Kurvenwertes um einen Betrag $> \text{Grubenlimit}$ definiert.

Das Zusammentreffen eines gültigen Maximums mit einem gültigen Minimum kennzeichnet den Übergang von einer Textzeile zu einem Zeilenzwischenraum. Zur genauen Bestimmung dieses Überganges dient der Parameter *Minimarand*.

Der Punkt in dem die Kurve zwischen einem gültigen Maximum und einem gültigen Minimum diese Schwelle schneidet, wird als Zeilenkante festgelegt.

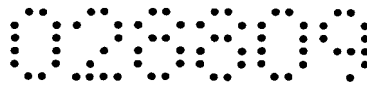
Zur Bestimmung des linken Randes einer Zeile wird ein Spaltenhistogramm gemäß

$$f(x) = \sum_{i=0}^{\text{Height}-1} (\text{BlackPixel})$$

x Spaltenindex des Bildausschnittes
Height Bild-Höhe

erstellt, in Worten die Farbinformation der Bildpunkte jeder Spalte des erfaßten Bildausschnittes wird aufsummiert. Der linke Textrand ist (bei Vorhandensein zumindest einer Zeile) durch einen sprunghaften Anstieg des Funktionswertes $f(x)$ definiert.

Die Nachführung der Zeilen, d.h. die Information über den weiteren Verlauf der Zeilen, die besonders bei



- 6 -

handgeführten Lesevorrichtungen aufgrund der bei diesen auftretenden Schwankungen von Bedeutung ist, erfolgt auf der Grundlage der Position der erkannten Zeichen.

Dazu werden die erkannten Zeichen in folgende Größengruppen eingeteilt:

Kleine Zeichen (z.B. 'a')	0,7*Zeilenhöhe
Große Zeichen (z.B. 'A', 'g')	Zeilenhöhe
Übergroße Zeichen (z.B. '[', 'j') (Unterlängen)	Zeilenhöhe+0,3*Zeilenhöhe

Sonderzeichen: die Zeichen lassen sich von der Größe nicht eindeutig zuordnen

Für die Ermittlung der neuen Textzeilenunterkante werden folgende Zeichengruppen unterschieden:

Baseline-Zeichen (z.B. 'A', '.'): die Zeichenunterkante entspricht der Textzeilenunterkante, unabhängig von der Größe des Zeichens;

Unterlängenzeichen (z.B. 'g', '['): die Zeichenunterkante entspricht der Unterlängengrenze, unabhängig von der Größe des Zeichens;

Sonderzeichen: diese Zeichen lassen sich bezüglich ihrer Unterkante nicht eindeutig zuordnen.

Auf der Grundlage dieser Zuordnungen und einem im Zuge des Klassifikationsverfahrens erhaltenen Wahrscheinlichkeitswert G betreffend die korrekte Klassifizierung des Zeichens wird nun die neue Zeilenhöhe $Height$ wie folgt ermittelt:

$G = \text{Wahrscheinlichkeit} * \text{CYC_MAX_WEIGHT}$

$$Height = \frac{\sum_{i=0}^{\text{CYC_MAX_EXTRPAR}-1} OldHeight[i] + NewHeight * G}{\text{CYC_MAX_EXTRPAR} + G}$$

G Gewichtung der aus dem aktuellen Zeichen abgeleiteten Zeilenhöhe
 Wahrscheinlichkeit Wahrscheinlichkeit der korrekten Zeichenklassifikation (Wertebereich zwischen 0 und 1)
 CYC_MAX_WEIGHT Maximale Gewichtung der neuen Zeichenposition (z.B.: 5)
 Height Nachgeführte Zeilenhöhe (Großbuchstabenhöhe)
 CYC_MAX_EXTRPAR Größe des Ringpuffers für die Mittelung (z.B.: 3)
 OldHeight[] Ringpuffer
 NewHeight Aus aktuellem Zeichen abgeleitete Zeilenhöhe (Großbuchstabenhöhe)
 i Index im Ringpuffer

Die Ermittlung des Verlaufs der Textzeilenunterkante erfolgt gemäß:

$$G = \left(\text{Wahrscheinlichkeit} + \frac{1}{\text{CYC_MAX_WEIGHT}} \right) * \text{CYC_MAX_WEIGHT}$$

$$\text{Increase} = \frac{\text{OldIncrease} + \text{NewIncrease} * G}{1 + G}$$

$$\text{Base} = \text{NewBase} + \frac{\text{Increase} * \text{DeltaX} + 50}{100}$$

G Gewichtung der neuen Zeichenposition
 Wahrscheinlichkeit Wahrscheinlichkeit der korrekten Zeichenklassifikation
 CYC_MAX_WEIGHT Maximale Gewichtung der neuen Zeichenposition (z.B.: 5)
 Increase Nachgeführte aktuelle Steigung der Baseline in %
 OldIncrease Vorherige Steigung der Baseline in %
 NewIncrease Aus Position des aktuellen Zeichens errechnete Steigung der Baseline in %

028809

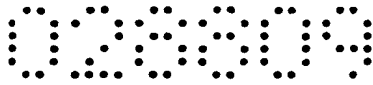
- 8 -

Base Nachgeführte Baselineposition (gerundet
auf ganzzahligen Wert)

NewBase Aus Position des aktuellen Zeichens
errechnete Baselineposition

DeltaX X-Abstand im Image der beiden
Mittelpunkte der zuletzt extrahierten Zeichen

Die 'Increase' wird durch das Plausibilitätslimit
CYC_MAX_ZEILENVERSATZ (im Pocket Reader: 15%) limitiert.



Patentansprüche

- 1) Verfahren zur Bestimmung der Lage von Textzeilen bei Texterkennungsaufgaben, **dadurch gekennzeichnet**, daß die Helligkeitsverteilung eines erfaßten Bildausschnittes entlang der Vertikale ermittelt wird und diese Helligkeitsverteilung gefiltert wird, daß Maximumwerte und Minimumwerte der solcherart erhaltenen Funktion ermittelt werden und auf der Grundlage dieser Extrema Schwellwerte berechnet werden die als Grundlage für die Unterscheidung zwischen Textzeile und Zeilenzwischenraum dienen.
- 2) Verfahren nach Anspruch 1, **dadurch gekennzeichnet**, daß ein Zeilenzwischenraum dann festgestellt wird, wenn die Funktion eine Kombination eines Maximums mit einem Minimum aufweist, bei der das Minimum einen Wert kleiner als $\text{Funktionsminimum} + \text{Anzahl der Bildpunkte über die Breite des Bildausschnittes}/15 + 2 \cdot \text{Anzahl der Bildpunkte über die Breite des Bildausschnittes}/15 \cdot \text{Funktionsmaximum}/\text{Anzahl der Bildpunkte über die Breite des Bildausschnittes}$ aufweist und der Abfall der Funktionswerte nach dem Maximum einen Wert größer $(\text{Funktionsmaximum} - \text{Funktionsminimum})/2$ aufweist.
- 3) Verfahren nach Anspruch 1 oder 2, **dadurch gekennzeichnet**, daß zur Feststellung des linken Rands der Textzeilen vor der Bestimmung der Lage der Textzeilen die Helligkeitsverteilung eines erfaßten Bildausschnittes entlang der Horizontale ermittelt wird und die solcherart erhaltene Funktion den Zeilenbeginn durch einen sprunghaften Anstieg des Funktionswertes darstellt.
- 4) Verfahren nach einem der Ansprüche 1 bis 3, **dadurch gekennzeichnet**, daß nach der erstmaligen Feststellung der Lage einer Zeile der weitere Verlauf derselben durch Auswertung der Information über die erkannten Textzeichen ermittelt wird.

008809

- 10 -

Zusammenfassung:

Es wird ein Verfahren zur Bestimmung der Lage von Textzeilen bei Texterkennungsaufgaben angegeben, bei dem die Helligkeitsverteilung eines erfaßten Bildausschnittes entlang der Vertikale ermittelt wird und diese Helligkeitsverteilung gefiltert wird, bei dem Maximumwerte und Minimumwerte der solcherart erhaltenen Funktion ermittelt werden und auf der Grundlage dieser Extrema Schwellwerte berechnet werden, die als Grundlage für die Unterscheidung zwischen Textzeile und Zeilenzwischenraum dienen.

Das Verfahren ist besonders vorteilhaft bei handgeführten elektronischen Lesevorrichtungen einsetzbar.

Siemens Aktiengesellschaft Österreich
durch



Handlungsvollmacht

Zl. 2275 /Präs. 96

GR 98P7016 AT

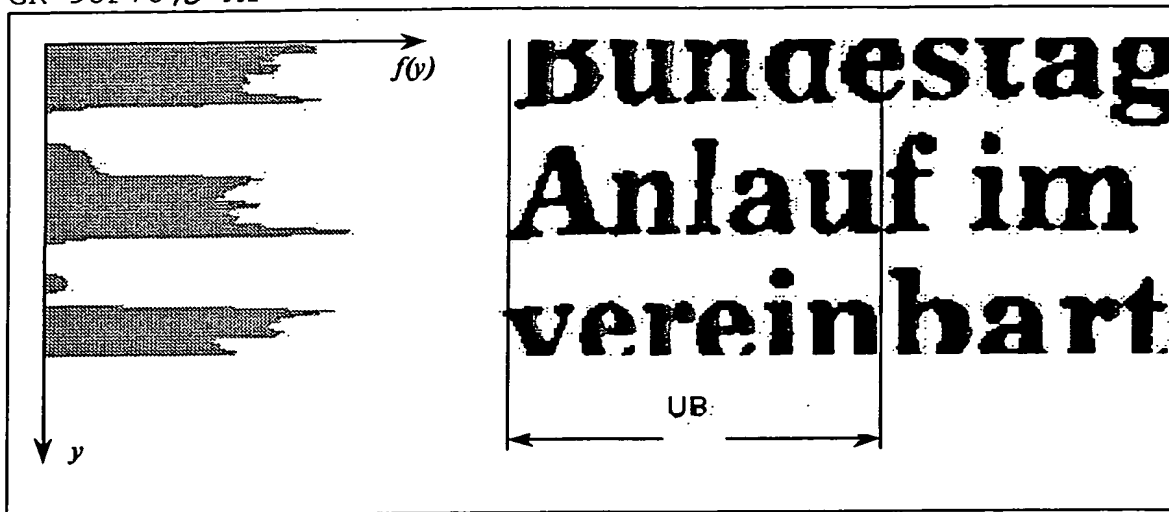


Fig. 1

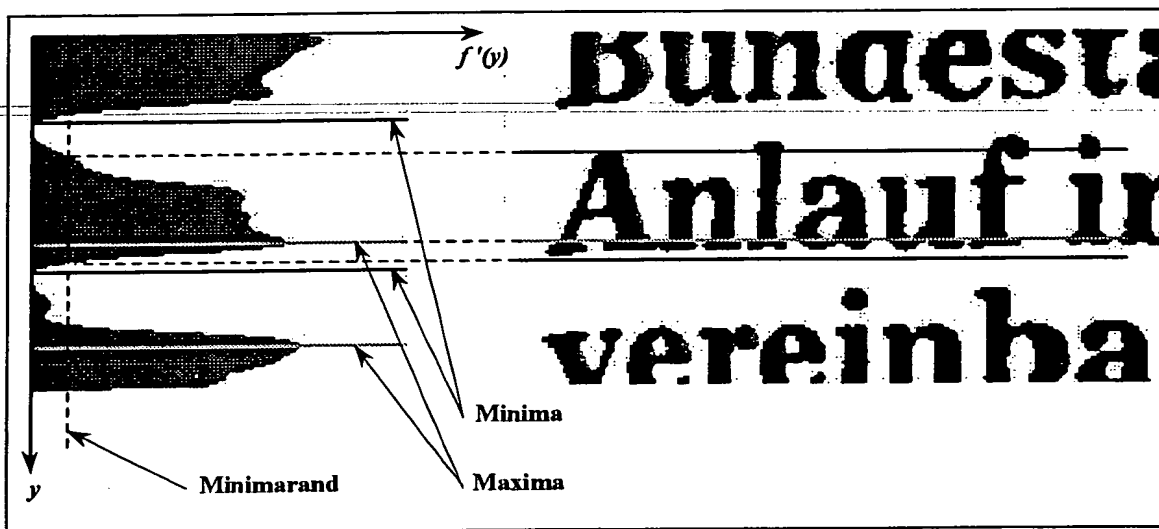


Fig. 2

Siemens Aktiengesellschaft Österreich

durch

Handlungsvollmacht

Zl. 2275 /Präs. 96

THIS PAGE BLANK (USPTO)